Environmental Health Perspectives
Vol. 53, pp. 185–187, 1983

Commentary

Additional Thoughts on the Review of Epidemiologic Studies Related to Ingested Asbestos

by Gary M. Marsh*

Following my presentation (*1*), a number of very interesting questions were raised. I am happy to be able to discuss some of these issues in this commentary. One question dealt with the fact that the concentration of fibers in Connecticut (*2,3*) was very low. Could any positive result really be expected? Since a dose-response relationship between ingested asbestos and related cancers has not been established, it is not possible to characterize any exposure level as insufficient for increasing cancer incidence or mortality above background levels. In fact, this basic uncertainty has probably provided much of the motivating force behind the ecological studies that have been conducted to date. If one subscribes, however, to the notion of the "existence" of a positive dose-response relationship between ingested asbestos and cancer, then, of course, the probability of detecting an elevated risk would be enhanced by studying populations that are exposed to higher levels of waterborne asbestos.

Another question concerned the use of multiple regression in the Connecticut study: How well did the data fit the normal assumptions for multiple regression? The questioner felt that a perfect fit cannot be expected, and that these kinds of data are usually fraught with outliers. He also suggested that the variance of observations is also probably not constant. I have found that it is not uncommon to be faced with interpreting the results of a statistical procedure in the absence of information regarding the adherence of the data to the assumptions underlying the procedure. The assumptions noted by the questioner (normality and equality of variances) are only two of several fundamental assumptions underlying the valid use of multiple regression analysis as a tool of statistical inference. It is likely in the Connecticut study that these other assumptions (e.g., negligible measurement error associated with independent variables and assumption of linearity) were also not completely satisfied. Fortunately, many parametric statistical procedures, such as multiple regression, have been shown by Monte Carlo studies and other techniques to be fairly robust (or perform at an adequate level) with respect to moderate deviations from the underlying assumptions. I suspect that the deviations from the underlying assumptions associated with the Connecticut study were not severe enough to have drastically altered the overall conclusions that were reported. The uncertainties regarding this issue could be obviated, however, if information regarding these important assumptions were provided in conjunction with study results.

One interesting issue that was raised concerns the fact that in the regression analysis of the San Francisco studies (*4,5*), if O/E was zero, the authors added a constant before taking logs. The dependent regression variable was then $\log [(O/E) + K]$. For the rarer tumors, many tracts would have $O/E = 0$ and all of these tracts would have $\log [(O/E) + K] = \log K$. Could this drastically affect the regression results?

It should first be pointed out that the transformation, $\log [(O/E) + K]$, was used by Kanarek et al. (*4*) for all values of O/E and not just when O/E was zero. The inappropriateness of the $\log [(O/E) + K]$ transformation was addressed by Tarter in 1981 (*6*). Tarter notes that this transformation can be written as: $\log [(O/E) + K] = \log [(O + EK)/E] = \log (O + EK) - \log (E)$. Tarter then reasons that since the argument of $\log (E)$ is

*Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15261.

supposed to predict O itself and not $(O + EK)$ it is better to use the variate $\log [(0 + K)/(E + K)]$ in place of $\log [(O/E) + K]$. Tarter also goes on to show that by choosing a large value for K (specifically, $K = 32$) in his form of the transformation (rather than $K = 0.01$ as used by Kanarek et al.) that the overall significance level of the asbestos in drinking water cancer association is greatly enhanced.

One question concerned my summary table of weaknesses and asked whether it was proper in a table such as this that "uncontrolled confounding" can contribute up to eight "points," while "ecologic study design" and other items can contribute only one "point"?

I constructed Table 4 with the express purpose of summarizing the various limitations/weakness inherent in the 13 reviewed studies. The aggregation of "points" (which represents the total number of limitations/weaknesses noted) shown at the bottom of the table for each study was done for purely descriptive purposes and in no manner whatsoever influenced the overall integration of the research findings. In more statistical terms the total "points" value was considered as a nominally scaled variable rather than as an ordinal or interval scaled measure. Because of this the total "points" was not used, for example, to rank the studies from 1 to 13 on the basis of their total limitations/weaknesses. If an objective comparative analysis such as ranking were attempted it would be improper, as the questioner noted, to aggregate "points" without weighting the individual points according to their relative importance or impact.

An important question concerned whether it was really valid to assume that $p = 0.05$ or 0.10 in my large deviation probability P_D calculations? As I had pointed out in my discussion of the site summary tables, my judgment of association was subjective. To convert that type of judgment into exact p values and then to come up with P_D values such as those in Table 5 was, the questioner felt, to give the informal subjective procedure more scientific weight than it had.

It is important first to distinguish clearly the purpose and reasoning behind the large deviation probability (P_D) calculations and the calculations made to assess the level of interstudy male-female agreement (ϕ coefficients with associated Fisher-Irwin test). The P_D calculations were performed only to more objectively evaluate the extent to which the number of male and female findings to date may be due to chance rather than biologic or other factors. Such an objective assessment is always important in problems involving simultaneous inference since a certain number of statistically significant comparisons can always be expected to occur due to chance factors alone. As footnoted in the text of my paper, at the 5% level of significance the probability of falsely claiming statistical significance in at least one of n independent comparisons is $1 - 0.95^n$. Thus, the P_D values related only to the extent that the absolute number of positive findings for males and females deviate from the expected number and do not represent the level of agreement between male and female findings *per se*. Moreover, the p values of 0.05 or 0.10 chosen for the P_D calculations are not objective characterizations of subjective judgments but rather represent only the theoretical probability of observing, in any individual study, a positive finding for males or females due to chance alone, that is, under the null hypothesis that no real biologic ingested asbestos-cancer association was present. The 5% and 10% levels are consistent with the conventional Type I statistical error rates under which most epidemiologic research is performed. The resulting P_D calculations then represent the overall probability of jointly observing in $n_{..}$ independent studies, due to chance factors alone, $n_{1.}$ or more and $n_{.1}$ or more positive associations in males and females, respectively. A $P_D < 0.05$, for example, implies that there is less than a 5% chance that the absolute number of observed positive male and female findings was due solely to chance factors given a 5% (or 10%) chance of a false positive finding in any one individual study. In other words, a small P_D value argues against the likelihood that a given number of positive male and positive female findings represents only chance phenomena. However, as mentioned in my text, the P_D approach did suffer from the limitations of small sample size ($n_{..}$) and the necessity to assume that the $n_{..}$ independent studies provided qualitatively and quantitatively equivalent information toward the integration of findings for any given cancer site. For these reasons it was recommended and adopted that the results of the P_D analysis should not be regarded as conclusive assessments of risk but rather should serve only as a rough guide for the direction and emphasis of future research. In this context, the p values which were of concern earlier become essentially arbitrary since it was primarily the relative relationship of the cancer site-specific P_D values rather than their absolute magnitude which ultimately generated the recommended priority of specific etiologic hypotheses to be explored in future research. The arbitrariness of the undelying p value is evident since this prioritization or

ranking based on P_D values is completely invariant to the choice of p.

Since the P_D values do not explicitly account for the overall degree of interstudy male-female agreement, ϕ coefficients and associated Fisher-Irwin probability levels were computed and interpreted in conjunction with the P_D values. Assuming that the biologic effects of ingested asbestos are similar in males and females (i.e., a no interaction model), a good level of agreement in male-female findings (a ϕ coefficient close to unity) also argues against the likelihood that a given pattern of observed findings represents only chance phenomena. In this sense, the P_D values and ϕ coefficients are complementary; that is, for a given cancer site a low P_D value in conjunction with a ϕ coefficient close to unity is more suggestive of a true ingested asbestos–cancer association than is a low P_D value and a ϕ coefficient close to zero. Unfortunately, because of very small sample sizes, ϕ coefficients for most of the cancer sites shown in Table 5 could not be computed.

Finally, I have been asked to comment on the utility of future ecologic studies on the topic of ingested asbestos, with or without multiple regression. As noted in the text of my paper, I agree that while the ability to make a causal inference from ecologic data often can be enhanced using more sophisticated analytic techniques, such as multiple regression, there will always remain an element of uncertainty until the etiologic hypotheses generated from ecologic studies are tested more definitively at the individual level rather than at the group level.

The views and policies presented by the author in this commentary do not necessarily reflect those of the U.S. Environmental Protection Agency. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

REFERENCES

1. Marsh, G. M. A critical review of epidemiologic studies related to ingested asbestos. Environ. Health Perspect. 53: 49–56 (1983).
2. Harrington, J. M., Craun, G., Meigs, J. W., Landrigan, P. J., Flannery, J. T., and Woodhull, R. S. An investigation of the use of asbestos cement pipe for public water supply and the incidence of gastrointestinal cancer in Connecticut, 1935-1973. Am. J. Epidemiol. 107: 96–103 (1978).
3. Meigs, J. W., Walter, S., Heston, J., Millette, J. R., Craum, G. F., Woodhull, R. S., and Flannery, J. T. Asbestos-cement pipe and cancer in Connecticut, 1955-1974. J. Environ. Res. 42: 187–191 (1980).
4. Kanarek, M. S., Conforti, P., Jackson, L., Cooper, R. C., and Murchio, J. C. Asbestos in drinking water and cancer incidence in the San Francisco Bay Area. Am. J. Epidemiol. 112: 54–72 (1980).
5. Conforti, P. M., Kanarek, J., Jackson, L., Cooper, R. C., and Murchio, J. C. Asbestos in drinking water and cancer in the San Francisco Bay area 1969-1974 incidence. J. Chronic Dis. 34: 211–224 (1981).
6. Tarter, M. E. Pattern recognition in the context of an asbestos cancer threshold study. Proceedings of the 13th Annual Symposium on the Interface: Computer Science and Statistics, Springer Verlag, in press.